DOCUMENT RESUME

ED 228 283 TM 830 170

AUTHOR Martin, Paula
TITLE Considerations and Recommendations for Evaluating Bilingual Education Programs.
PUB DATE Nov 82
NOTE 28p.
AVAILABLE FROM California Evaluation Services, 2090 Green Street, No. 27, San Francisco, CA, 94123 ($10.00)
PUB TYPE Guides - Non-Classroom Use (055)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Accountability; *Bilingual Education Programs; Check Lists; Data Collection; *Evaluation Criteria; Evaluation Needs; Federal State Relationship; Feedback; *Information Needs; *Program Evaluation; State Standards

ABSTRACT

General considerations in evaluating bilingual education programs and data aggregation at the state and national levels are presented. Federal and state information requirements and accountability guidelines include the needs for information/data for various audiences, for evaluation resources, for increased project capability to conduct evaluations, for development of interest in evaluations, and for special studies. Evaluation standards for bilingual education programs consider (1) the evaluation model's characteristics; (2) its information coverage; (3) its general design characteristics; (4) its specific design characteristics; (5) data collection types and techniques; (6) analytic procedures; and (7) the evaluation's utility. An evaluation checklist of context and process variables to consider when conducting a bilingual program evaluation is presented. Specific criteria for report writing are discussed and a data analysis checklist with sample forms for displaying student achievement data is included. Useful suggestions for evaluation report writing and design developement are given. (PN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED228283

CONSIDERATIONS AND RECOMMENDATIONS

FOR

EVALUATING BILINGUAL EDUCATION PROGRAMS

November 1982

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

P. Martin

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paula Martin, Ed.D.
California Evaluation Services
2090 Green Street No. 27
San Francisco, CA 94123
(415) 922-6528

TM 830 170

© Paula Martin 1983

TABLE OF CONTENTS

I. INTRODUCTION. 1

II. GENERAL CONSIDERATIONS IN EVALUATING BILINGUAL PROGRAMS 2

Federal and State Considerations for Bilingual Evaluation . . 2
Table I: Bilingual Evaluation Issues and Needs 5
CES Evaluation Standards for Bilingual Education Programs . . 6

III. GENERAL EVALUATION RECOMMENDATIONS FOR BILINGUAL EDUCATION EVALUATION REPORTS. 9

CES Evaluation Checklist. 11
Data Analysis Checklist 18
Tables III,IV,V -Tables for Student Achievement Data. 19

IV. CONCLUSIONS 22

I. INTRODUCTION

This paper will present considerations and recommendations for evaluating bilingual education programs and data aggregation at a state or national level.

There are special considerations and issues in evaluating a bilingual program. Bilingual education programs, instructional models, and methodologies run a wide range of possibilities in program design. Evaluation strategies have to be adaptable to various program designs and should attempt to gather data regarding overall program outcomes and the programmatic variables, or specific instructional models, that caused those outcomes.

Students in bilingual education programs are impacted by other categorical aide programs like Title I and state monies. It is difficult to isolate the various effects of these instructional programs and additional support when a multiple funding/multiple treatment effect leads to many confounding variables in program treatment.

Another concern in bilingual program evaluations has been appropriate assessment instruments for limited-English-speakers. There is also a shortage of qualified evaluators who are bilingual program specialists as well.

There is a great variation in evaluation designs and reporting formats in bilingual program evaluations and this has limited possibilities for data aggregation and making any generalizations about program effect. There is a current need at local, state and federal levels for concrete information on the effectiveness of bilingual education. Without this effectiveness data, funds, legislation and support for bilingual education will continue to diminish.

This paper is an extension of the research conducted by the author for her doctoral dissertation at the University of San Francisco. The title

of the dissertation is:

"Evaluation in Bilingual Education
A Meta-Evaluation of Title VII Projects
1979-80"

A more standardized reporting format is needed which would be realistic in terms of the resources available to bilingual programs for conducting data collection through program evaluations.

II. GENERAL CONSIDERATIONS IN EVALUATING BILINGUAL PROGRAMS

Federal and State Level Considerations for Bilingual Evaluations

At a federal level, it would be most useful if data collection was more standardized regarding Title VII programs in terms of the type of information gathered to describe projects and the testing instruments used for student achievement. Standardized evaluation procedures and reporting formats, forms, questionnaires, and suggestions for data analysis would allow for more valid data comparisons at a state or national level.

Appropriate budgeting for evaluations must be encouraged to allow for sufficient evaluation resources. Federal interest, encouragement and guidelines for evaluation costs could make a difference with technical assistance and a standardized reporting format in a walk-through evaluation handbook.

Meta-evaluation feedback to districts from both a state and federal level regarding the quality of their evaluation reports would lead to an improvement in overall data collection. There could also be incentive rewards for districts which comply with evaluation criteria and guidelines and who conscientiously submit their reprots. Evaluations of poor quality could be rejected. Funds could be pulled for poor accountability.

Follow-up on designated projects could also be done through the Evaluation,

Dissemination, and Assessment Centers regarding special federal interests. There could be RFPs (Requests for Proposal) for controlled evaluation studies.

State level monitoring could invovle technical assistance workshops and assistance in implementing prescribed evaluation plans.

There needs to be clarification on state and federal information needs, regulations and guidelines. Evaluation requirements could be clearer to districts and request that at a minimum the reports included information regarding:

1. Evaluation intentions.
2. General program description and approach to bilingual education.
3. Number of students by language group, grade and language proficiency.
4. Staffing.
5. Major component objectives.
6. Pre-/post-test dates.
7. Names and descriptions of tests and test levels.
8. Student achievement data by language proficiency and grade.
9. Interpretation of results.
10. Conclusions.
11. Recommendations for planning and improvement.

Districts should be required to report raw scores by grade level and language proficiency and the numbers of students at various grade levels on suggested standardized tests to move data toward a form that could be more easily aggregated at a state or national level.

Dissemination of evaluation materials should control the number, types and quality of evaluation guidebooks and handbooks that are developed by state or federally funded contracts and available to Title VII projects regarding the bilingual program evaluations.

There is still a need for a state or federal comprehensive plan for bilingual project evaluations that outlines specific evaluation standards and options, in terms of models, designs, and data gathering instruments. Many districts do not have the degree of sophistication needed to develop appropriate instruments, and think in terms of research evaluation designs

for bilingual programs within the framework of their other evaluation needs and accountability needs for the variety of state and federal programs they may have.

Table I outlines various evaluation issues and needs at a local, state and federal level.

The following standards on pages 6 to 8 are suggested as a guideline for developing a comprehensive evaluation design for a bilingual education program. Student achievement data alone is only a part of the total evaluation plan.

BILINGUAL EVALUATION ISSUES AND NEEDS

TABLE I

Evaluation Issues and Needs	Solution
I. The Need for Information/Data for Various Audiences A. Local Students/Parents/Teachers/ Principals/Board Members B. State C. Federal	1. Various Reports for Different Audiences
II. The Need for Evaluation Resources A. Personnel: Evaluators, Assistants, Clerical Assistance B. Funding for the Evaluation C. Additional Consulting or Technical Assistance	1. Appropriate budgeting 2. Federal interest, encouragement, and guidelines for evaluation costs 3. Graduate students 4. Evaluation, Dissemination, and Assessment Centers (EDACs); Lau Centers; and State Departments
III. The Need to Increase Project Capability to Conduct Evaluations	1. Clarification on state and federal information needs, regulations, and guidelines 2. Technical assistance 3. Standardized reporting format, checklists, questionnaires, and scoring formats in a walk-through evaluation handbook approach
IV. The Need to Develop Interest in Evaluations	1. Incentives (positive and negative) at a state and federal level 2. Meta-evaluation feedback (state and federal) 3. Federal pulling of funds or rejection of evaluations for poor quality
V. The Need for Special Studies	1. Request for proposals for specially-funded contracts to conduct controlled evaluations on selected projects

P. Martin
Evaluation Consultant
CES 10/80

CES EVALUATION STANDARDS FOR BILINGUAL EDUCATION PROGRAMS

A complete evaluation design should consider information regarding the following:

I. Evaluation Model Characteristics:
   1. Type of Model
      A. Descriptive
      B. Goal-based
      C. Objective-based
      D. Formative/Summative
      E. Specific theoretical model (e.g., CIPP)
   2. Funds for the Evaluation
   3. Evaluation Resources
   4. Staffing for the Evaluation

II. Information Coverage (Scope of the Report)
   1. Total Program
   2. Selected Components - Which?
   3. Student Achievement Data--Which areas, subjects?
   4. Cost-Information
   5. Definition of the Audiences for the Evaluation

III. General Design Characteristics
   1. Purpose of the Evaluation
   2. Evaluation Objectives
   3. Evaluation Timeline
   4. Discussion of Limitations/Constraints
   5. Discrepancy of intended objectives and observed performance
   6. Discussion of cause and effect factors
   7. Allowance for noting unanticipated outcomes
   8. Current Needs Assessment Data
   9. Possible special studies and longitudinal studies

©1981 Paula Martin

IV. Specific Design Characteristics

1. Testing schedule and population included
2. Tests to be Used - Standardized - Criterion Referenced
3. Norms Used for Reference
4. Control or Comparison Groups
5. Sampling Techniques
6. Entry/Exit Criteria
7. Other standards, levels of minimum competency, continuums of skills, etc.

V. Data Collection Types and Techniques

1. Questionnaires - To Whom?
2. Interviews (Telephone or Face to Face) - To Whom?
3. Reports from Project Staff
4. Special Data Collection
   (a) suspensions/disciplinary records/dropout rates
   (b) attendance patterns (students)
   (c) attendance patterns, logs, calendars of activities and other records
5. Special Evaluation Activities
   (a) use of evaluation teams
   (b) coordination with State or Federal program review efforts
   (c) use of parents/community in evaluation

VI. Analytic Procedures

1. Type of Statistics
   (a) means
   (b) standard deviations
   (c) percentiles
   (d) t-test
   (e) regression techniques
2. Special Approaches
   (a) longitudinal analysis
   (b) special comparison studies/groups
3. Data Processing Techniques
   (a) SPSS
   (b) BIOMED

4. Quantified Ratings from Observations/Questionnaires
5. Documentation and No Statistical Procedures
6. Summaries and narrative reports
7. Other Analysis
8. Validity and Reliability Issues

VII. Evaluation Utility

1. Schedule of reports and type of reports
2. Audiences for evaluation information
3. Dissemination of evaluation results
4. Use of results for planning, decision-making, program information.

III. GENERAL EVALUATION RECOMMENDATIONS FOR BILINGUAL EDUCATION EVALUATION REPORTS

Some practical considerations regarding evaluation reports, report formats, general data collection, and analysis will be presented here. An overall checklist for bilingual program evaluations which include various context and process variables is included on pages 11 to 14. Allowances have to be made regarding the range of possibilities in program implementation and program design and the resources available for conducting evaluations. These recommendations will have to be modified to fit the project's individual information needs at a local, state and federal level. The intention here is to offer some general guidelines to report writing and evaluation design.

Report Cover/Construction

The best report cover is velobind with plastic ribs and a clear plastic cover so that the project title and project/grant number show through. If a velobind is not available, the report should be securely bound with a stitch staple (2) or other strong binding. The report should be able to be easily identified. The project year(s) should be on the cover and evident. The name of the evaluator(s) who prepared the report should be on the title page. Pages should be numbered. Typing should be professional and duplicated copies should be clear.

General Report Format

The general report format should include a cover page, title page, table of contents, the body of the report, and appendices when needed.

Abstract

A two-page abstract should be included in the beginning of the report that states the funding level, overall district enrollment, number of students

in the Title VII project, grade levels, and major evaluation findings.

Introduction

There should be an introduction to the evaluation report that covers the purpose of the report, the evaluation objectives, evaluation questions to be answered, and gives a recapitulation of the evaluation design.

Program Characteristics

The report will be descriptive to a certain extent in regarding the characteristics of the Title VII project. The Checklist on pages 11 to 14 lists more specific context and process variables to consider in describing the Title VII project and evaluation efforts. The context of the program should be described in terms of geographic location and the characteristics of the school district.

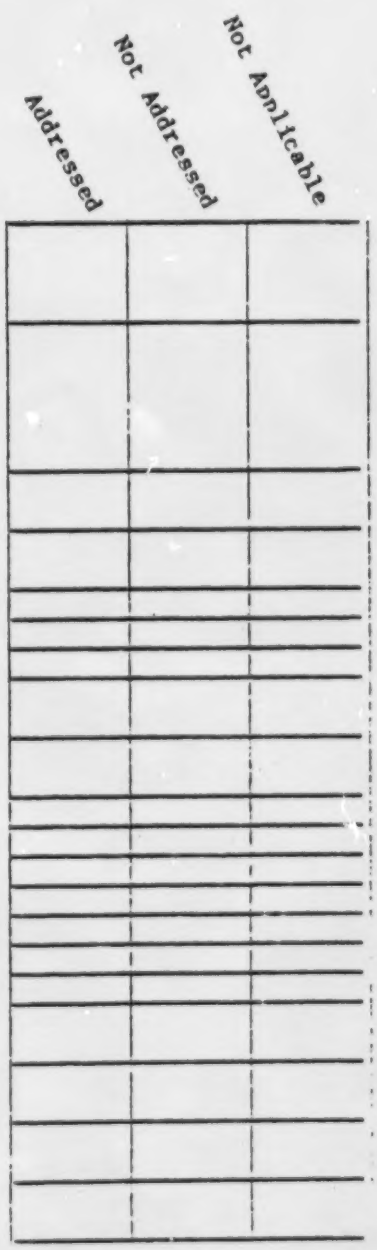
The student population of the school district by ethnicity and grade should be presented on a table to put the Title VII project in a clearer context. The district count on non-English-proficient, limited-English-proficient students should be presented as well. Following, one could give an overview of Title VII student enrollment data by grade, language proficiency and ethnicity, and other characteristics to give a clearer picture of the student population in the bilingual program.

The historical background of the bilingual program and changes in emphasis or grade levels should be mentioned. The years of previous funding and current funding level are important and should be clear.

Staffing for the project should be presented graphically with an indication of the numbers of teachers, paraprofessionals, resource teachers, directors, and the time they spend with the programs. The number of credentialled teachers working with the program and those who are on waivers, etc., could be presented as well. The degree of bilinguality among the staff and their

CES EVALUATION CHECKLIST

Context and Process Variables to Consider in Conducting a Bilingual Program Evaluation



I. Context of the Program

A. Geographic Location

B. Description of the District
Population by Ethnicity and Grade
LEP/NEP Count
Socioeconomic Factors

C. Available Resources

District
County
State
Other

D. Community Characteristics

Ethnic composition
Socioeconomic factors
Migration
Housing
Services: health and welfare
General occupations
Local politics
Local resources

II. Program Characteristics

A. Size of the program/Project

B. Funding Level

C. Historical Background

© Paula Martin 1981

	Addressed	Not Addressed	Not Applicable
D. Student Characteristics			
Ethnicity			
Age			
Previous educational background			
Aspirations/attitudes			
Abilities in the native language (L_1) and second language (L_2)			
Transiency			
Absenteeism			
E. Staffing (Teachers, Paraprofessionals, and Others)			
Time with Title VII Project			
Type of teaching credential			
# of years of teaching/classroom experience			
# of years in bilingual education			
Degree of bilinguality			
Staff utilization			
F. Implementation Plan			
Goals and objectives			
Proposed activities			
Timeline for activities			
G. Program Design			
Self-contained bilingual classroom			
Pull-out bilingual instructor/ESL instructor			
Departmentalized			
Alternate day			
Other			
H. Academic Program			
Entry and exit criteria			
Average hours of language usage in L_1 and L_2			

© Paula Martin 1981

	Addressed	Not Addressed	Not Applicable
Types of material in use			
Curriculum			
I. Instruction			
1. Grouping (average hours per week)			
Whole class instruction			
Individualized instruction			
2. (Average Hours) Time allocated per student per week			
Average teacher hours			
Average adult hours			
3. Average English as a Second Language (hours per week)			
Number of groups for English			
Average number in group			
4. Teaching Methodologies			
Preview/Review			
Concurrent			
J. Parent Participation/Community Involvement			
Roles of parents and community			
Parent education			
Volunteer efforts			
Impact of parent participation/ community involvement			
K. Staff Development			
Number of hours inservice for staff and others			
Types and topics of inservice programs			

© Paula Martin 1981

	Addressed	Not Addressed	Not Applicable
Value of programs			
Application of knowledge acquired to classroom situation			
III. Evaluation Characteristics			
Funds for evaluation and other resources			
Type of evaluator/contractor			
Timeline for the evaluation			
Provisions for assessing student achievement and program impact			
Structure for a longitudinal following of student progress			
On-going utility of the evaluation process			

© Paula Martin 1981

former teaching experiences or experience in bilingual education are other variables that could be considered.

The implementation plan for the program in terms of goals, objectives, proposed activities, and the timeline for implementation should all be considered. The instructional setting including the program design (self-contained classrooms, pull-out, etc.) and instructional methodologies (preview-review, concurrent, etc.), curriculum, and materials in use are important. More details on the academic program could involve the balance of language usage in the classroom in the native language (L_1) and second language (L_2), time allotted per student each week for various types of instruction in large or small groups or on a one-to-one basis.

Other components of the bilingual program should be described and evaluated including parent participation/community involvement and staff development. The evaluation component should be described in terms of the resources allocated for evaluation efforts, the type of evaluator/contractor and proposed activities for data collection and the use of evaluation results.

Reporting Evaluation Findings

Survey questionnaire data, student achievement data, and general program monitoring can all make up the general findings upon which conclusions can be based.

Survey Questionnaire Data

A survey questionnaire to teachers, paraprofessionals, and other staff, covering the effectiveness of the overall program and various aspects of the program is valuable to identify what they see as strengths or weaknesses.

A survey questionnaire to parents and community workers regarding the value of various activities and their perception of the strengths and weaknesses of the program also adds to the evaluation.

An inservice evaluation form shoudl also be used for each inservice session, and all forms can be summarized at the end of the year for an overview as to how valuable the staff development program was to various participants.

Student questionnaires are very appropriate for bilingual programs, especially at the secondary level regarding scheduling, counseling, and the general curriculum. Student questionnaires also can explore the affective domain to determine students' self-esteem, students' aspirations, etc.

Student Achievement Data

Student achievement data should be reported by language proficiency, grade level, and the numbers of students in various subgroups by grade level. If possible, scores should be considered according to instructional treatment and other context and process variables.

The dates of pre/post-testing should be clearly stated on each display of student achievement data. Testing instruments should be described in enough detail to make the name and level of the test obvious.

Appropriate assessment instruments are a primary concern in assessing language proficiency or general abilities in the first and second languages utilized in the bilingual program.

Standardized tests should have norms that include students somewhat similar to the target population (i.e., language, author, socioeconomic status, etc.). Criterion-referenced tests should be described in terms of their validity or reliability.

Projects should avoid the 12 common hazards in evaluation as listed in "A Practical Guide to Measuring Project Impact on Student Achievement" (Horst, Tallmadge and Wood, 1975), namely:

1. The use of grade-equivalent scores.

2. The use of gain scores.

3. The use of norm-group comparisons with inappropriate test dates.

4. The use of inappropriate levels of tests.

5. The lack of pre/post-test scores for each project participant

6. The use of non-comparable treatment and comparison groups.

7. The selection of project participants based on pre-test scores.

8. The assembling of a matched comparison group after the the project participants are selected.

9. The careless administration or scoring of tests.

10. The assumption that an achievement gain is due to the greatment when, in reality, it could be due to some other factor.

11. The use of non-comparable pre-tests and post-tests.

12. The use of inappropriate formalae to estimate post-test scores.

A Data Analysis Checklist developed by the RMC Research Corporation, as a result of their Project Information Packages contract, delineates suggested steps for data analysis for bilingual program evaluations. (See Table II.)

Suggested layouts for the display of student achievement data are included in Tables III, IV, and V which follow. These tables ate revised forms of formats used in 1979-80 Title VII reports and would display raw scores, percentiles, normal curve equivalent stanine (NCEs), and grade-equivalent scores. These tables may be an example of displays of data that would be understandable and useful to several audiences although technically grade-equivalent scores are not recommended.

There should be a description of the data analysis or scoring format

Table II

Data Analysis Checklist
Outline

I. General principles

A. Analyze data both by individual years for short-term goals and cumulatively for long-term goals.

B. Separate data according to language proficiency groups.

C. Separate data further according to instructional treatment.

II. Preparation (applies to most evaluation designs)

A. Convert raw scores to standard scores (preferably normalized standard scores such as NCEs). Use these scores for all analyses.

B. Separate out those students with both pre- and posttests.
1. Compute means and standard deviations.
2. Plot the distributions of pretest scores.
3. Plot the distributions of posttest scores.
4. Plot the joint distribution of pretest and posttest scores.

C. For students with pretest scores only:
1. Compute the mean and standard deviation.
2. Plot the distribution of scores.

D. For students with posttest scores only.

Save the scores for student files and for use as next year's pretest scores.

III. Check for irregularities in the data:

A. Floor or ceiling effects.

B. Large changes in standard deviations from pretest to posttest.

C. Low correlations between pre- and posttest scores, or irregular joint distributions.

D. Differences between students who took the posttest, and those who dropped out.

E. Look for any other features of the data that strike you as strange, and be sure that you can explain them. Ideally, item data should be examined.

IV. Apply the statistical or other procedures relevant to the particular evaluation design in use.

Be sure that your analyses are relevant to the questions you are trying to answer.

Taken from the RMC Research Corporation Report UR 460 "An Evaluation of the Project Information Packages (PIPs) As Used for the Diffusion of Bilingual Projects," Vol. 3, May 1980, p. 165.

Table III

Table for Student Achievement Data

School District _____ Group Characteristics _____

School _____ ___ Pullout ___ Bilingual S/C ___ Other

Grade _____ MES/NEP LES/LEP FES/FEP Comparison

Grade	READING											
	N	RS	$\bar{X}_{RS}$	Md_{RS}	$\bar{X}_{p}$	Md_{p}	$\bar{X}_{NCE}$	Md_{NCE}	$\bar{X}_{s}$	Md_{s}	$\bar{X}_{GE}$	Md_{GE}
1st												
2nd												
3rd												
4th												
5th												
6th												

N = Number of Students
RS = Raw Score
$\bar{X}_{RS}$ = Mean Raw Score
Md_{RS} = Median Raw Score

$\bar{X}_{p}$ = Mean Percentile
Md_{p} = Median Percentile
$\bar{X}_{NCE}$ = Mean (NCE) Normal Curve Equivalent
Md_{NCE} = Median (NCE) Normal Curve Equivalent

$\bar{X}_{s}$ = Mean Stanine
Md_{s} = Median Stanine
$\bar{X}_{GE}$ = Mean Grade Equivalent Score
Md_{GE} = Median Grade Equivalent Score

Tests Used

	READING					MATH				
	Date	Name	Level	Form	Subtest	Date	Name	Level	Form	Subtest
PRE										
POST										

© Paula Martin 1981

Paula Martin, Ed.D.
Evaluation Consultant
2090 Green Street, No. 27
San Francisco, CA 94123

Table IV

Table for Student Achievement Data

School District _____

School _____

Grade _____

Group Characteristics

____ Pullout ____ Bilingual S/C ____ Other

NES/NEP LES/LEP FES/FEP Comparison

	Non English		Limited English		Fluent English		English Dominant		
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	
N									
$\bar{X}$									
Md									
σ									
$\bar{X}_p$									
Md_p									
$\bar{X}_{NCE}$									
Md_{NCE}									
$\bar{X}_{GE}$									
Md_{GE}									

N = Number of Students
$\bar{X}$ = Average Raw Score
Md = Median Raw Score
σ = Standard Deviation of Raw Score

$\bar{X}_p$ = Mean Percentile
Md_p = Median Percentile
$\bar{X}_{NCE}$ = Mean (NCE) Normal Curve Equivalent

Md_{NCE} = Median (NCE) Normal Curve Equivalent
$\bar{X}_{GE}$ = Mean Grade Equivalent Score
Md_{GE} = Median Grade Equivalent Score

Tests Used

	READING					MATH				
	Date	Name	Level	Form	Subtest	Date	Name	Level	Form	Subtest
PRE										
POST										

Suggestions for Reporting Student Achievement Data

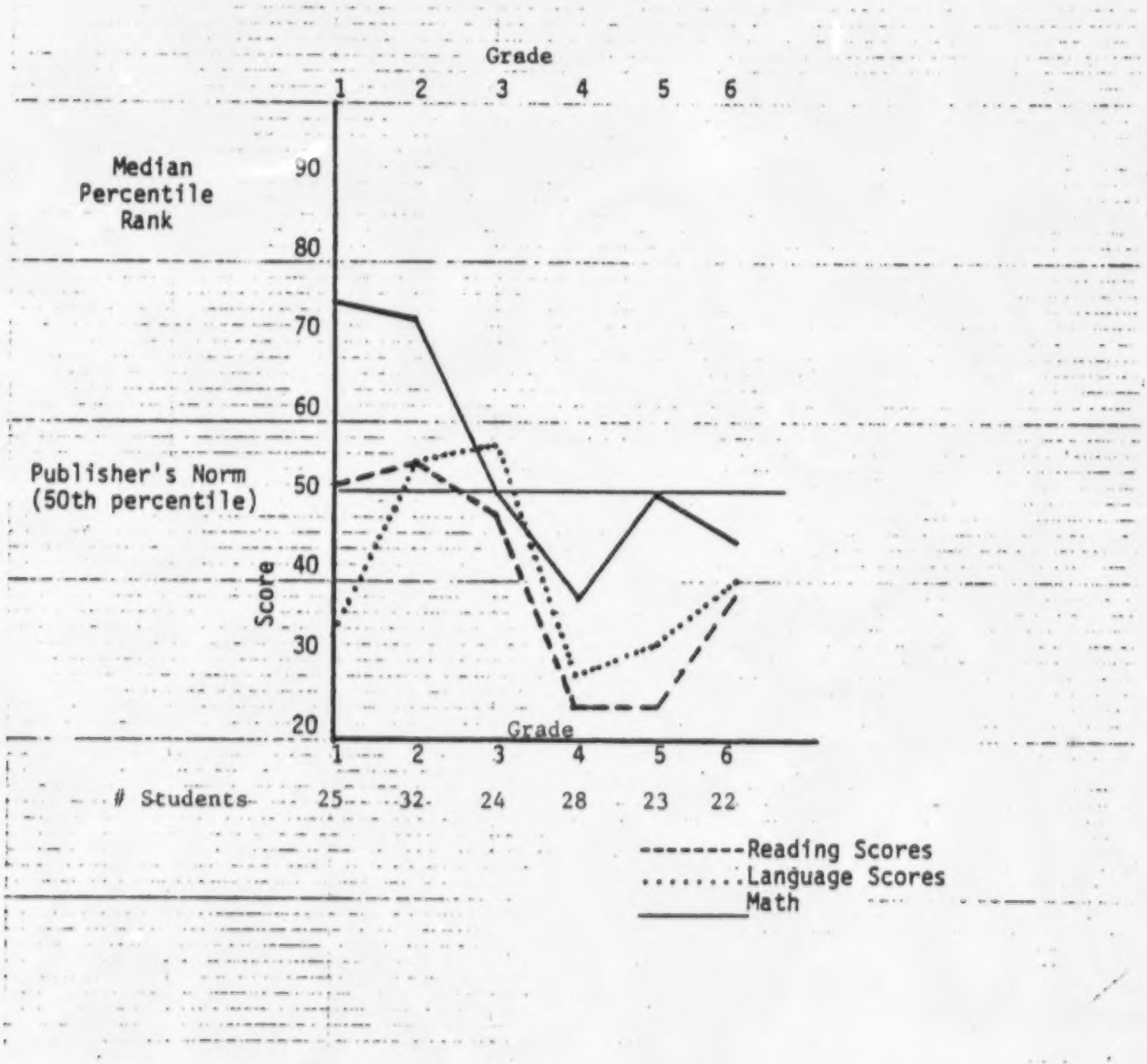
Student achievement data should be presented by grade and language proficiency and the number of students in each subgroup. A complete interpretation and discussion of results should be included.

Evaluation reports can present a breakdown of achievement results in terms of raw scores, percentiles, stanines, and/or NCEs. The use of grade-equivalent scores is not recommended. Means and medians of each category can be examined. The median raw score or percentile, etc., is more appropriate than the mean in examining bilingual test scores since there may be a wide range of achievement scores due to the background of the students. The mean is more sensitive to extreme scores and may give a distorted notion as to what is considered "average." The distance between the mean and the median may be examined for indications about the shape of the distribution of scores. If the mean is smaller than the median, the distribution is negatively skewed indicating a group of higher scores. If the median is smaller than the mean, the distribution may be positively skewed indicating a grouping of lower scores.

The use of these suggested tables gives the evaluator or researcher the option of using various findings for different audiences or the possible aggregation of data from several projects.

Table V
Suggested Format for Displaying Student Schievement Data
(Percentiles Only)

This format for displaying student achievement data utilizes percentiles. Here reading, language, and math scores for grades 1-6 can be displayed in relation to the publishers norm of 50th percentile which is used as a cut-off score in many programs for compensatory education.



The advantage of this format is that it is understandable to parents and others who want to get an overview of student achievement at various grade-levels.

This data indicates that second grade students were above the 50th percentile in all subject areas while sixth, fifth, and fourth graders were below in all areas. There is also a fourth grade "slump" that is evident, perhaps in this case due to the change from elementary school to middle schools.

used in measuring student achievement, and there should be a substantial interpretation of the findings in a manner that is intelligible to a range of audiences (local, state and federal).

Student achievement data should be maintained on student profiles for each program participant by student number, and other characteristics. A computerized database should be developed where possible to maintain records for a possible longitudinal study. The use of micro-computers may be a possibility for districts where larger computer systems are not available. A data management system should be established.

General Program Monitoring

General program monitoring involves data collection on all aspects of the program and the various activities that took place over the funding period.

There are a range of records on students that are relevant to program effects including absenteeism, mobility, and suspension or drop-out rates. These should be considered in determining project impact.

Inservice activities and staff development activities should be documented by the numbers and roles of participants that attended. Dates for the sessions should be given and topics presented.

Parent/community activities can be documented by reporting the numbers of parents, community members, and staff attending various functions. The number of home visits and contacts made through the community liaisons and teachers should also be documented.

Visitors to program classrooms, and program review results should be part of an ongoing data collection process. An activity log or calendar of events with the numbers of participants at various functions should be kept that accounts for program activities over the past year.

Conclusions and Recommendations

Conclusions should provide a good overall interpretation of all of the evaluation findings, and recommendations should be strong and specific enough to be of some use to program planners and decision-makers.

Evaluation Utility

Evaluation utility can only be ensured if it is tied into the evaluation plan in an ongoing manner. Evaluation findings should be reported to several audiences and the respondents over the evaluation period. Results should be presented in various forms so that they are understandable to parents/community, staff, and others.

IV. CONCLUSIONS

In conclusion, there is a continued need to address the issue of a lack of national and state level data on the effects of bilingual education programs.

It is evident from research (Alkin, 1976; Burt and Dulay, 1979; Zappert and Cruz, 1977; California State Department of Education, 1979; and Martin, 1981) that there is a need to improve the quality of data collected annually through Title VII program evaluations.

There are possibilities for aggregating data on a state-wide or federal level through the use of evaluation reports. Government and state education agencies have only to be clear on their intentions for data gathering and offer federally-funded projects more specifics and guidance in meeting evaluation requirements. The lack of valid and reliable data is due to an uncertainty on the behalf of local education agencies in the amount of funds that should be put into evaluation efforts and what data they should be collecting. Standardized evaluation formats, questionnaires, and specific directions for

collecting, analyzing, and presenting student achievement data are necessary. This could be done through a handbook approach backed up with technical assistance and an evident interest in evaluations from funding agencies as reflected through meta-evaluation feedback and incentives (positive or negative) to encourage LEAs to do appropriate data collection.

This paper has addressed these issues and presented some general recommendations on conducting bilingual evalution and writing reports.